

IMPLEMENTASI ARSITEKTUR MEDALLION OPEN-SOURCE UNTUK ANALISIS KOMPARATIF TECH LAYOFFS GLOBAL

Abdulhadi Muntashir*¹, Firdaus Satrio Utomo¹, Rendy Ilyasa¹, Gladys Andromeda Bilqis¹, Mohamad Hilman¹

¹Program Studi Informatika, Universitas Sultan Ageng Tirtayasa, Indonesia

*e-mail: 3337230041@untirta.ac.id

Abstrak

Ekosistem *startup* Indonesia yang bergantung pada pendanaan ventura asing rentan terhadap gelombang *tech layoffs* global, tetapi analisisnya kerap terhambat inkonsistensi data geospasial dan biaya infrastruktur *cloud* yang tinggi. Penelitian ini bertujuan menunjukkan bahwa kombinasi *open-source* PostgreSQL dan *Data Build Tool* (DBT) berbasis Arsitektur Medallion mampu menghasilkan analisis komparatif PHK teknologi Indonesia dan global secara akurat tanpa platform berbayar. Metode ELT (*Extract, Load, Transform*) diterapkan dengan novelty berupa strategi *compound join key* untuk mengeliminasi anomali duplikasi baris (*Cartesian product*) akibat ketidakseragaman koordinat geospasial. Sistem berhasil mengeksekusi delapan model data dalam 1,19 detik dengan tingkat keberhasilan uji kualitas 100% (24/24 pengujian) dan *zero row inflation*. Hasil analisis mengungkap karakteristik spesifik PHK di Indonesia berupa dominasi sektor *Transportation* sebesar 76,81% dari total 2.721 pekerja yang tersentralisasi di Jakarta, berbeda signifikan dari pola global yang terdiversifikasi pada sektor *Consumer* dan *Retail*. Penelitian ini menunjukkan bahwa infrastruktur *open-source* berbasis PostgreSQL dan DBT mampu menghasilkan performa analitik yang andal dan efisien untuk mendukung analisis komparatif PHK teknologi serta mengidentifikasi konsentrasi kerentanan ekosistem digital nasional pada wilayah tertentu.

Kata kunci: arsitektur medallion; *compound join key*; *data warehouse*; DBT; *tech layoffs*.

Abstract

Indonesia's startup ecosystem, heavily reliant on foreign venture capital, is vulnerable to global tech layoffs waves, yet comparative analysis is frequently hindered by geospatial data inconsistencies and high cloud infrastructure costs. This study aims to demonstrate that an open-source combination of PostgreSQL and Data Build Tool (DBT) under a Medallion Architecture can produce accurate comparative analyses of technology layoffs in Indonesia and globally without relying on paid platforms. An ELT (*Extract, Load, Transform*) approach was implemented with a compound join key strategy as the primary novelty, specifically designed to eliminate row duplication anomalies (*Cartesian product*) caused by non-uniform geospatial coordinates. The system successfully executed eight data models in 1.19 seconds with a 100% data quality test pass rate (24/24 tests) and zero row inflation. Analysis results reveal a distinctive layoff pattern in Indonesia, with the Transportation sector dominating at 76.81% of 2,721 total affected workers concentrated in Jakarta, contrasting sharply with the more diversified global pattern led by the Consumer and Retail sectors. This research demonstrates that an open-source infrastructure based on PostgreSQL and DBT is capable of supporting reliable and efficient analytical processing for comparative tech layoffs analysis while revealing the concentration of vulnerabilities within Indonesia's digital ecosystem.

Keywords: *compound join key*; *data warehouse*; DBT; *medallion architecture*; *tech layoffs*.

1. PENDAHULUAN

Persaingan pasar yang semakin dinamis menjadikan pengambilan keputusan berbasis data (*data-driven decision-making*) sebagai aspek fundamental bagi organisasi dalam mempertahankan performa dan pertumbuhan [1]. Kebutuhan ini mendorong investasi pada infrastruktur *scalable data warehouse* guna menangani volume dan variasi data yang terus meningkat [2]. Salah satu pendekatan modern yang digunakan untuk menjaga integritas informasi adalah Arsitektur Medallion, yang menerapkan prinsip penyempurnaan data progresif melalui lapisan pemrosesan yang sistematis [3]. Implementasi arsitektur ini semakin optimal dengan adopsi metodologi *Extract, Load, Transform* (ELT) melalui perangkat *Data Build Tool* (DBT) yang memungkinkan transformasi data dilakukan secara langsung di dalam *data warehouse* dengan standar rekayasa perangkat lunak yang tinggi [4]. Fenomena pemutusan hubungan kerja massal (*tech layoffs*) di industri teknologi global telah bertransformasi dari krisis sementara di awal pandemi menjadi pola pengurangan tenaga kerja yang terus berlanjut hingga awal 2026. Data dari *Layoffs.fyi*, gelombang PHK mencapai titik tertinggi pada kuartal I 2023 dengan lebih dari 167.674 karyawan terdampak secara global. Meskipun volume tahunan sempat menurun pada periode 2024-2025, terjadi peningkatan jumlah PHK yang signifikan pada awal tahun, seperti kenaikan dari 1.552 orang pada Desember 2025 menjadi 24.928 orang pada Januari 2026

[5]. Penelitian sebelumnya yang menggunakan model epidemiologi SIR memprediksi intensitas PHK akan kembali normal pada akhir 2023 [6]. Namun, fakta empiris memperlihatkan bahwa proses ini berlangsung lebih lama dan kompleks sehingga data PHK teknologi memiliki pola yang dinamis dan memerlukan infrastruktur data yang mampu mendukung proses integrasi serta analisis secara akurat.

Perusahaan teknologi di Indonesia, seperti GoTo Group, Xendit, dan ByteDance, juga mengalami pengurangan tenaga kerja dalam jumlah signifikan selama periode 2022–2024. Kondisi ini menunjukkan bahwa ekosistem *startup* Indonesia yang bergantung tinggi pada pendanaan ventura asing memiliki tingkat kerentanan yang tinggi. Dominasi sektor *Fintech* dan *Transportation* di Indonesia menjadikan analisis komparatif dengan tren global relevan untuk memahami pola PHK pada ekosistem *startup* nasional. Proses analisis komparatif ini sering terhambat oleh kendala teknis rekayasa data, seperti inkonsistensi penamaan sektor industri dan potensi duplikasi baris (*Cartesian product*) pada operasi *join* data geospasial. Arsitektur *data warehouse* tradisional dinilai kurang efisien karena adanya risiko latensi data [7]. Implementasi Arsitektur Medallion melalui pendekatan ELT memang terbukti mampu menjaga integritas data, tetapi penerapannya masih didominasi platform *cloud* berbayar seperti Databricks dan Snowflake yang berisiko menimbulkan biaya operasional tinggi serta *vendor lock-in* [8], [9]. Platform *cloud fully-managed* juga membatasi fleksibilitas konfigurasi dan optimasi *pipeline* data geospasial. Keterbatasan ini menjadi krusial ketika menangani data geospasial yang tidak terstruktur, di mana fleksibilitas optimasi langsung pada mesin basis data diperlukan untuk meningkatkan akurasi *join* pada koordinat yang tidak konsisten.

Penelitian oleh Hartono *et al.* [10] telah memetakan perbedaan karakteristik sektor terdampak antara skala global dan regional Indonesia, tetapi analisis tersebut masih terbatas pada pendekatan deskriptif tanpa dukungan infrastruktur data yang kuat. Sementara itu, Hutabalian *et al.* [11] menunjukkan bahwa otomatisasi *pipeline* data berbasis PostgreSQL mampu mempercepat proses analisis, tetapi belum mengintegrasikan strategi penanganan anomali geospasial secara eksplisit. Keterbatasan dari kedua penelitian tersebut menjadi celah penelitian (*research gap*) yang mendasari motivasi penelitian ini, yaitu kebutuhan akan solusi *open-source* yang tidak hanya efisien secara komputasi, tetapi juga mampu menangani inkonsistensi koordinat geospasial melalui strategi kunci komposit yang terverifikasi secara kuantitatif. Tantangan teknis tersebut menjadi semakin relevan pada konteks Indonesia, di mana data geospasial perusahaan teknologi lokal kerap mengalami inkonsistensi format koordinat dan penamaan kota yang tidak seragam antar-sumber. Kondisi ini secara langsung memotivasi pengembangan strategi *compound join key* dalam penelitian ini sebagai solusi yang meningkatkan akurasi penggabungan data lokasi pada ekosistem *startup* Indonesia yang terpusat di Jakarta. Sebagai landasan teoretis, berbagai literatur terdahulu telah mengeksplorasi pilar teknologi rekayasa data yang relevan dengan permasalahan tersebut. Penerapan Arsitektur Medallion memiliki peran krusial dalam menjaga integritas transaksi dan adaptabilitas skema pada data berskala besar [12]. Penelitian terdahulu di bidang data spasial juga banyak mengandalkan kapabilitas PostGIS yang mampu memproses data lokasi jauh lebih cepat dibandingkan sistem konvensional [13]. Sementara itu, pemantauan indikator kinerja secara multidimensi umumnya bertumpu pada desain sistem berbasis OLAP [14]. Aspek kualitas dan konsistensi pada *pipeline* data juga biasanya diperkuat oleh prinsip *Continuous Integration* (CI) yang terbukti mampu meningkatkan keandalan *deployment* secara keseluruhan [15].

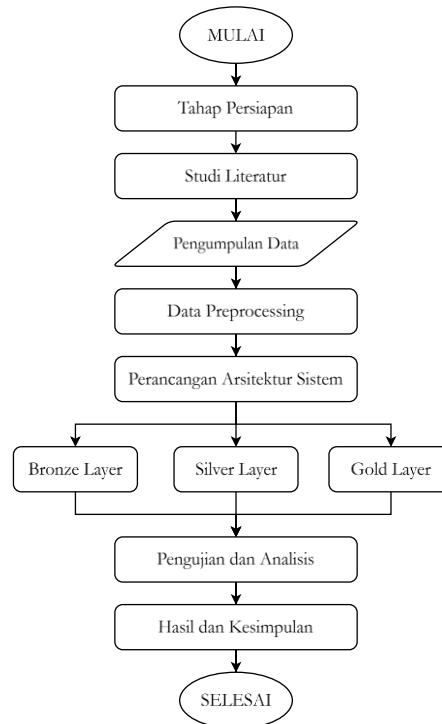
Penelitian ini mengusulkan implementasi Arsitektur Medallion secara mandiri menggunakan PostgreSQL dan DBT untuk mengolah 1.745 insiden *layoffs* dari 1.374 perusahaan di 58 sektor industri. *Novelty* penelitian ini terletak pada penerapan *compound join key* berbasis atribut geospasial di dalam lingkungan DBT untuk mengatasi anomali *Cartesian product* pada proses *join* data lokasi yang tidak konsisten. Strategi ini mengombinasikan atribut kota, negara, *latitude*, dan *longitude* sebagai kunci komposit dalam operasi *join* antar tabel sehingga mampu meningkatkan akurasi integrasi data spasial dibandingkan pendekatan *join* konvensional berbasis atribut tunggal. Tujuan utama penelitian ini adalah mengevaluasi implementasi kombinasi alat *open-source* dalam mendukung performa analitik yang efisien untuk analisis PHK teknologi secara akurat. Kontribusi penelitian ini diharapkan bisa menjadi referensi implementasi *data warehouse* berbasis *open-source* untuk analisis data geospasial dan multidimensi pada ekosistem *startup* digital. Struktur naskah ini disusun ke dalam lima bagian utama, yaitu Pendahuluan, Metode, Hasil Penelitian, Pembahasan, dan Kesimpulan.

2. METODE

Penelitian yang berkonsep rekayasa data (*data engineering*) menggunakan metode deskriptif serta implementasi untuk membangun *data warehouse* berbasis *Medallion Architecture* melalui analisis dari data

pemutusan hubungan kerja (PHK) dari perusahaan teknologi tahun 2020-2025. Alasan menerapkan ini karena mampu mendukung proses integrasi, transformasi, serta analisis data secara terstruktur sehingga dapat menghasilkan informasi yang akurat dan mudah dipahami.

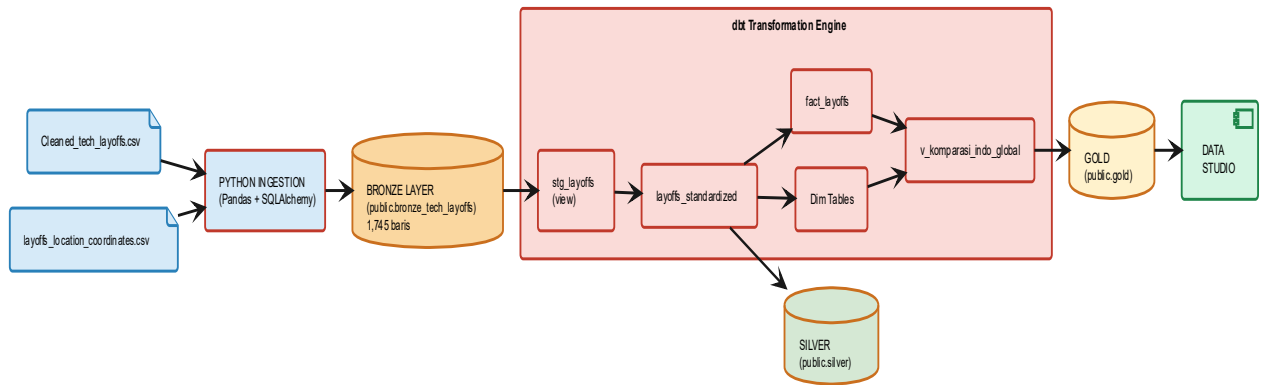
Dataset diambil melalui platform Kaggle. Dataset berupa informasi perusahaan seperti jumlah karyawan, lokasi perusahaan, industri, serta waktu terjadinya PHK. Data yang berhasil didapat selanjutnya dilakukan analisis untuk memastikan kelengkapan, konsistensi, dan kesesuaian format dengan yang lain sebelum dilakukan proses pengolahan data. Proses ini dapat dilihat melalui Gambar 1.



Gambar 1. Alur Penelitian.

Alur penelitian ini diawali dengan tahap persiapan sebagaimana ditunjukkan pada Gambar 1, yang meliputi pendefinisian parameter sistem serta batasan masalah yang akan dikaji. Selanjutnya, proses inti penelitian dimulai dengan pengumpulan data sekunder berupa dataset “Tech Layoffs 2020-2025” yang diperoleh dari platform Kaggle. Dataset ini merupakan hasil integrasi dari dua sumber utama, yaitu layoffs.fyi dan peerlist.io, yang mendokumentasikan kasus pemutusan hubungan kerja (PHK) di sektor teknologi secara global. Dataset tersebut didistribusikan di bawah lisensi terbuka CC BY-NC-SA 4.0. Tahap berikutnya adalah *data preprocessing*, yang mencakup pembersihan data untuk menangani nilai kosong, penghapusan duplikasi, serta penyeragaman format penulisan. Proses ini sangat penting untuk memastikan kualitas data sebelum memasuki tahap selanjutnya. Perancangan arsitektur sistem menggunakan pendekatan Arsitektur Medallion dilakukan segera setelah data memenuhi standar kebersihan.

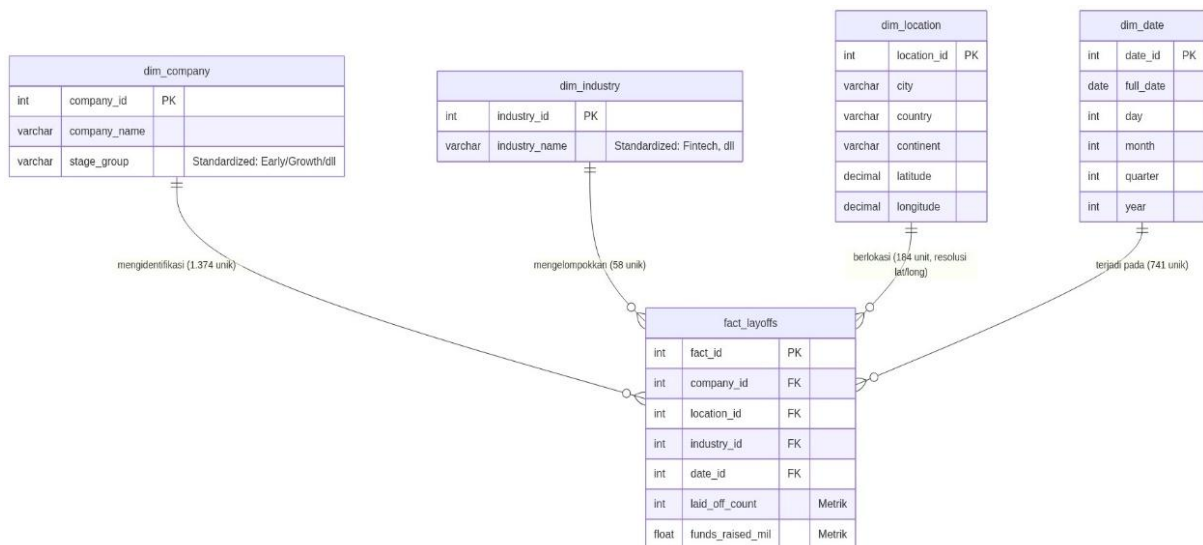
Arsitektur Medallion membagi proses pengolahan data berdasarkan tingkat kematangannya. Lapisan *Bronze Layer* berfungsi menyimpan data mentah sebagaimana adanya untuk menjaga integritas sumber historis. Tahap berikutnya dilakukan pada *Silver Layer* berupa transformasi dan standardisasi sehingga data menjadi lebih bersih, konsisten, dan siap dianalisis. Lapisan akhir yaitu *Gold Layer* menyusun data ke dalam model dimensional yang telah dioptimalkan untuk mendukung analisis dan visualisasi pada *dashboard*. Rincian implementasi *pipeline* arsitektur ini ditampilkan pada Gambar 2. Penelitian ini diakhiri dengan tahap pengujian dan analisis untuk memvalidasi kinerja sistem, serta penarikan kesimpulan berdasarkan temuan yang diperoleh.



Gambar 2. Diagram Arsitektur Medallion Pipeline.

Proses *ELT (Extract, Load, Transform)* dilakukan menggunakan skrip Python dengan bantuan pustaka *Pandas* dan *SQLAlchemy* untuk melakukan ingest data mentah dari file CSV ke dalam sistem. *PostgreSQL* dipilih karena memiliki skalabilitas *query* yang baik. Seluruh transformasi data dikelola menggunakan *DBT* yang bekerja secara bertahap di dalam *database* untuk memastikan integritas dan kualitas informasi di setiap lapisan Arsitektur Medallion.

Model dimensional yang diimplementasikan menerapkan pendekatan *Star Schema* yang terdiri dari tabel fakta dan tabel dimensi. Tabel fakta (*fact_layoffs*) menyimpan metrik utama berupa volume PHK, sementara tabel dimensi menyimpan atribut deskriptif seperti profil perusahaan, industri, lokasi geografis, dan waktu kejadian. Struktur ini dirancang untuk mengoptimalkan performa *query* analitik serta memudahkan integrasi data menuju alat visualisasi. Arsitektur teknis *pipeline* tersebut ditunjukkan secara detail pada Gambar 3.



Gambar 3. Entity Relationship Diagram (ERD) Star Schema.

Visualisasi data menggunakan *dashboard* analitik menjadi tahap berikutnya yang dilakukan setelah proses pembangunan *data warehouse* selesai. *Dashboard* digunakan untuk menampilkan informasi terkait tren PHK berdasarkan waktu, industri, dan lokasi perusahaan sehingga mempermudah pengguna dalam melakukan analisis data. Selanjutnya, evaluasi sistem dilakukan secara kuantitatif menggunakan tiga metrik utama, yaitu: (1) *row inflation rate* untuk mengukur efektivitas *compound join key*, (2) *data quality pass rate* untuk mengukur keandalan *pipeline*, dan (3) waktu eksekusi *pipeline* sebagai indikator performa sistem. Evaluasi ini bertujuan untuk menguji konsistensi data hasil transformasi serta memastikan bahwa data yang dihasilkan sesuai dengan sumber data awal. Dengan demikian, sistem *data warehouse* yang dibangun diharapkan mampu mendukung proses analisis data PHK perusahaan teknologi secara lebih efektif dan terstruktur.

3. HASIL PENELITIAN

A. Implementasi dan Evaluasi *Pipeline* ELT

Implementasi *pipeline* ELT dimulai dengan proses ingesti terhadap 1.745 insiden PHK teknologi dari 1.374 perusahaan di 58 sektor industri menggunakan skrip Python dengan bantuan pustaka Pandas dan SQLAlchemy. Seluruh data mentah berhasil dimuat ke dalam *Bronze Layer* pada PostgreSQL tanpa kesalahan. Selanjutnya, proses transformasi dijalankan oleh DBT yang mengeksekusi 8 model data, meliputi tahap *staging*, normalisasi empat tabel dimensi (perusahaan, waktu, industri, lokasi), satu tabel fakta, serta satu *view* agregasi komparatif. Validasi integritas data dilakukan bersamaan dengan proses transformasi melalui mekanisme pengujian otomatis DBT *tests*, verifikasi *row parity*, dan evaluasi strategi *compound join key*. Pengujian otomatis mencakup 24 skenario uji yang memvalidasi aspek *not null*, *uniqueness*, integritas referensial, serta pengujian kustom *assert_valid_coordinates* untuk validitas data geospasial. Ringkasan performa eksekusi dan hasil validasi seluruh tahapan *pipeline* dirangkum dalam Tabel 1.

Tabel 1. Performa Eksekusi dan Validasi Kualitas Pipeline.

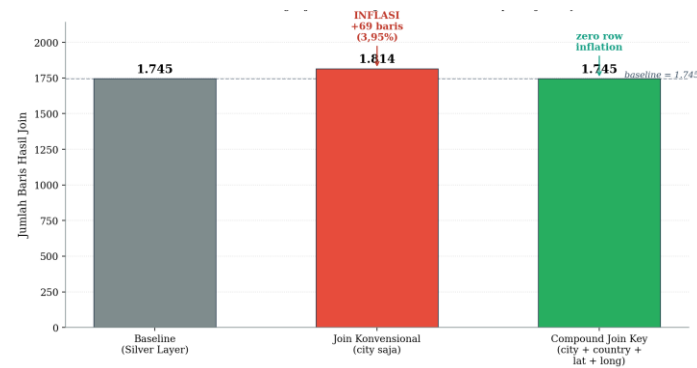
Tahapan <i>Pipeline</i>	Metrik Utama	Durasi	Status
<i>Ingestion</i> (Python)	1.745 baris	< 0,01 detik	Sukses
DBT <i>Transformation</i>	8 model	1,19 detik	Sukses
<i>Data Quality Test</i>	24 skenario	1,33 detik	24/24 lulus

Verifikasi *row parity* dilakukan untuk menguji efektivitas strategi *compound join key* dalam mengeliminasi anomali duplikasi baris (*Cartesian product*). Pendekatan *join* konvensional yang hanya menggunakan satu atribut lokasi (misalnya nama kota) berpotensi menghasilkan inflasi baris akibat ketidakseragaman data koordinat. Strategi *compound join key* mengatasi hal ini dengan mengombinasikan empat atribut secara simultan (kota, negara, *latitude*, dan *longitude*) sehingga setiap operasi penggabungan hanya menghasilkan pasangan data yang akurat. Tabel 2 menampilkan distribusi jumlah baris pada setiap tabel yang dihasilkan oleh *pipeline*.

Tabel 2. Distribusi Jumlah Baris pada Setiap Layer Pipeline.

Lapisan	Nama Tabel	Jumlah Baris
<i>Bronze</i>	Data mentah (sumber)	1.745
<i>Gold</i>	Tabel fakta	1.745
<i>Gold</i>	Dimensi perusahaan	1.374
<i>Gold</i>	Dimensi industri	58
<i>Gold</i>	Dimensi lokasi	184
<i>Gold</i>	Dimensi waktu	741
<i>Gold</i>	<i>View</i> komparasi Indonesia–Global	396

Jumlah baris pada tabel fakta di *Gold Layer* yang tetap berjumlah 1.745 identik dengan data sumber pada *Bronze Layer* mengonfirmasi tercapainya kondisi *zero row inflation*. Perbandingan kuantitatif antara kedua strategi *join* ditampilkan pada Gambar 4. Pendekatan *join* konvensional yang hanya menggunakan atribut kota menghasilkan 1.814 baris (inflasi sebesar 3,95% atau 69 baris), sedangkan *compound join key* menghasilkan tepat 1.745 baris tanpa inflasi sama sekali. Hasil ini menjadi bukti langsung bahwa *compound join key* berhasil mengeliminasi anomali *Cartesian product* pada seluruh operasi penggabungan data yang melibatkan atribut geospasial.



Gambar 4. Perbandingan Jumlah Baris: Join Konvensional vs Compound Join Key.

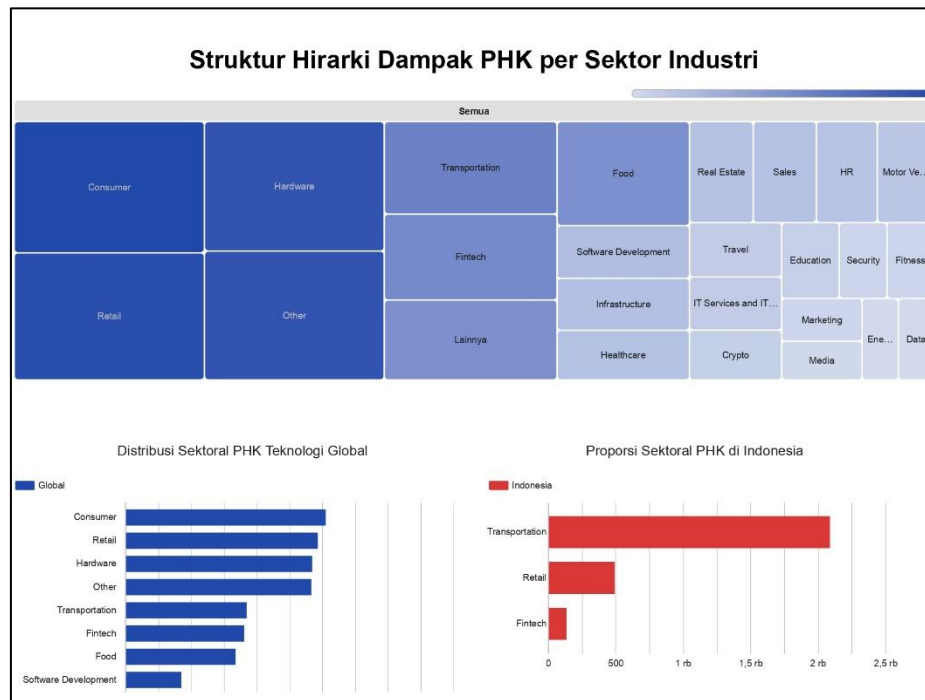
B. Analisis Komparatif Multidimensi PHK Indonesia dan Global

Analisis komparatif dampak PHK teknologi antara Indonesia dan skala global dilakukan pada tiga dimensi, yaitu temporal, sektoral, dan spasial yang seluruhnya terintegrasi dalam sistem *Business Intelligence*. Gambar 5 menyajikan analisis tren temporal serta dampak korporasi yang memperlihatkan perbandingan volume PHK tahunan dan peringkat perusahaan dengan dampak terbesar.



Gambar 5. Dashboard Tren Temporal dan Analisis Dampak Korporasi PHK (2020–2025).

Visualisasi pada Gambar 5 menunjukkan bahwa puncak PHK global secara temporal tercatat pada tahun 2023 dengan 176.616 pekerja terdampak. Sementara itu, guncangan terbesar di Indonesia terjadi lebih awal, yaitu pada tahun 2022 dengan total 1.931 pekerja terdampak. Dashboard tersebut juga mengidentifikasi perusahaan besar seperti Intel, Amazon, dan Microsoft dari sisi korporasi secara global. Data pada Gold Layer memperlihatkan fenomena PHK tetap berlanjut hingga awal 2026. Tren temporal ini juga menunjukkan lonjakan anomali secara konsisten pada setiap kuartal pertama. Perbedaan pola ini kemudian dianalisis lebih lanjut pada dimensi sektoral untuk melihat industri mana yang paling rentan, sebagaimana ditampilkan pada Gambar 6.



Gambar 6. Dashboard Analisis Hirarki dan Proporsi Sektoral PHK.

Visualisasi pada Gambar 6 menunjukkan adanya dominasi sektor tertentu pada masing-masing skala. Dampak PHK secara global tersebar luas dengan sektor Consumer (61.155 pekerja) dan Retail (59.228 pekerja) sebagai kontributor terbesar. Sementara di Indonesia, sektor Transportation menjadi kontributor tunggal terbesar dengan total 2.090 pekerja, diikuti Retail (494) dan Fintech (137). View komparasi Indonesia-Global pada Gold Layer menunjukkan konsentrasi pada sektor Transportation di Indonesia berbanding terbalik dengan distribusi global yang lebih merata tanpa dominasi tunggal. Terakhir, dimensi spasial dipetakan untuk melihat konsentrasi geografis sebaran insiden tersebut sebagaimana ditampilkan pada Gambar 7.



Gambar 7. Dashboard Analisis Spasial-Sektoral dan Peta Sebaran PHK Global.

Peta sebaran pada Gambar 7 menunjukkan konsentrasi yang tinggi di wilayah Amerika Utara dan Eropa. Detail koordinat yang bersifat teknis untuk wilayah Indonesia disajikan secara formal pada Tabel 3

guna menjamin akurasi posisi geografis yang dianalisis dalam sistem, meskipun visualisasinya menunjukkan titik persebaran.

Tabel 3. Konsentrasi Spasial-Sektoral PHK Teknologi di Indonesia.

Kota	Negara	Latitude	Longitude	Sektor Dominan	Total PHK
Jakarta	Indonesia	-6,175	106,827	Transportation	2.090

Temuan ini menunjukkan adanya konsentrasi PHK pada wilayah tertentu yang berpotensi memengaruhi ketahanan ekosistem digital regional. Konsentrasi ekstrem pada satu titik koordinat Jakarta (-6,175; 106,827) merupakan bukti validasi teknis sistem yang hanya bisa terdeteksi secara akurat berkat strategi *compound join key*. Ketidakteraturan koordinat pada pendekatan *join* konvensional berpotensi memecah data Jakarta menjadi beberapa entitas berbeda, sehingga dominasi sektor *Transportation* sebesar 76,81% bisa saja tidak teridentifikasi atau terdistorsi. Sistem yang dibangun menunjukkan bahwa infrastruktur *open-source* berbasis PostgreSQL dan DBT mampu mengungkap karakteristik spasial yang spesifik ini secara andal sebagai kapabilitas yang krusial bagi pengambilan kebijakan ketenagakerjaan digital di tingkat nasional.

4. DISKUSI

Implementasi Arsitektur Medallion berbasis teknologi *open-source* membuktikan kemampuannya mendukung analisis data secara akurat dan efisien. Analisis temporal memperlihatkan perbedaan pola gelombang PHK teknologi antara skala global dan Indonesia. Perbedaan pola ini mengindikasikan tingkat kerentanan ekosistem *startup* Indonesia terhadap pengetatan likuiditas global. Ekosistem tersebut merespons kondisi pasar melalui langkah efisiensi tenaga kerja lebih awal. Temuan ini mengonfirmasi penelitian Mayer [16] mengenai perilaku perusahaan yang memiliki tingkat eksposur tinggi terhadap pendanaan eksternal. Perusahaan dengan karakteristik ini cenderung memprioritaskan koreksi tenaga kerja lebih awal sebagai strategi manajemen risiko likuiditas. Evaluasi tingkat sektoral menyoroati dominasi sektor *Transportation* di Indonesia. Dominasi ini membuktikan sentralisasi kerentanan ekonomi digital nasional pada model bisnis *ride-bailing* dan logistik berbiaya operasional tinggi. Pola ini mengontraskan tren global yang mendistribusikan dampak PHK lebih luas ke sektor *Consumer* dan *Retail*. Perbedaan karakteristik tersebut menginterpretasikan kondisi kerentanan struktural ekonomi digital nasional yang belum mencapai tingkat diversifikasi industri secara merata. Temuan ini memperdalam analisis Hartono *et al.* [10] yang sebelumnya mengidentifikasi perbedaan sektoral secara deskriptif, dengan membuktikan secara kuantitatif bahwa ketimpangan tersebut berpusat pada model bisnis berbiaya operasional tinggi.

Penelitian ini mengonfirmasi studi Gujjala [12] mengenai peningkatan integritas transaksi melalui pembagian lapisan data pada Arsitektur Medallion. Sistem yang dievaluasi memperlihatkan keunggulan efisiensi infrastruktur menggunakan basis data dan perangkat transformasi *open-source* tanpa bergantung pada layanan berbayar. Keandalan infrastruktur analitik ini memfasilitasi penarikan wawasan temporal yang memberikan perspektif baru terhadap dinamika industri teknologi. Bukti kelanjutan tren PHK membantah proyeksi model epidemiologi SIR yang memprediksi normalisasi intensitas pemangkas tenaga kerja pada akhir tahun 2023 [6]. Penolakan terhadap model prediktif tersebut justru memperkuat validitas teori *overhiring* Mayer [16]. Mayer mengategorikan rentetan PHK massal sebagai mekanisme koreksi pasar pasca-ekspansi berlebihan selama masa pandemi. Pola ini dikonfirmasi oleh data Indonesia pada *Gold Layer* yang mencatat angka PHK mendekati nol sepanjang tahun 2021, bertepatan dengan puncak fase ekspansi digital ekosistem *startup* nasional (Gambar 5). Penelitian ini mengisi celah keilmuan dengan mengintegrasikan kapabilitas rekayasa data terapan dan analisis sosiogeografis industri.

Sistem yang diajukan menawarkan implikasi praktis bagi organisasi dalam merancang *open-source data warehouse* tanpa mengorbankan kualitas informasi. Strategi *compound join key* menunjukkan efektivitas tinggi dalam mengeliminasi anomali *Cartesian product* saat mengintegrasikan data geospasial. Efektivitas ini memvalidasi prinsip dependensi fungsional pada desain basis data relasional [17]. Kombinasi atribut *latitude* dan *longitude* mengakselerasi proses disambiguasi lokasi dengan nomenklatur serupa di berbagai wilayah geografis. Pendekatan kunci komposit ini menjamin tingkat akurasi integrasi data yang melampaui metode penggabungan atribut tunggal [13]. Penerapan infrastruktur analitik mengidentifikasi pemusatan spasial ekstrem pada wilayah ibu kota. Kondisi *Jakarta-centric* menyoroati tingginya risiko ketenagakerjaan regional akibat konsentrasi aktivitas ekonomi digital. Pemangku kebijakan membutuhkan analisis spasial ini sebagai

landasan penyusunan strategi pemerataan ekosistem teknologi nasional. Penelitian ini juga memperluas landasan teoretis penerapan *Continuous Integration* (CI) melalui penggunaan modul pengujian terotomatisasi. Sistem penjaminan mutu ini mengonfirmasi hipotesis Garudasu *et al.* [15] terkait peran CI dalam meningkatkan kualitas peluncuran rekayasa data secara komprehensif. Batasan cakupan pangkalan data mewajibkan interpretasi objektif terhadap temuan penelitian ini. Sumber data agregasi cenderung merekam insiden di perusahaan berskala menengah dan besar. Kecenderungan ini menciptakan potensi pelaporan yang lebih rendah (*underreporting*) bagi kasus PHK pada ekosistem perusahaan rintisan skala kecil dan UMKM teknologi di Indonesia. Sistem saat ini juga belum mengintegrasikan data milik institusi pemerintah, sehingga membatasi komprehensivitas analisis dampak sosial berskala makro. Kekurangan ini memberikan ruang eksplorasi lanjutan bagi pengembangan arsitektur *data warehouse* yang lebih inklusif.

5. KESIMPULAN

Penelitian ini berhasil mengimplementasikan Arsitektur Medallion berbasis PostgreSQL dan DBT untuk mendukung analisis komparatif *tech layoffs* Indonesia dan global. *Pipeline* yang dibangun mampu berjalan secara efisien dengan waktu eksekusi 1,19 detik serta tingkat keberhasilan pengujian kualitas data sebesar 100%. Penerapan strategi *compound join key* terbukti efektif dalam mengatasi anomali *Cartesian product* pada integrasi data geospasial dengan menghasilkan *zero row inflation* pada tabel hasil transformasi. Hasil analisis juga menunjukkan adanya perbedaan karakteristik PHK teknologi antara Indonesia dan skala global. Di Indonesia, PHK didominasi oleh sektor *Transportation* sebesar 76,81% yang terkonsentrasi di Jakarta, sedangkan pola global menunjukkan distribusi yang lebih beragam pada sektor *Consumer* dan *Retail*. Temuan ini mengindikasikan bahwa ekosistem digital Indonesia masih memiliki tingkat ketergantungan yang tinggi terhadap sektor dan wilayah tertentu. Secara praktis, penelitian ini menunjukkan bahwa pendekatan *open-source* berbasis PostgreSQL dan DBT dapat menjadi alternatif yang efisien dan fleksibel dalam pembangunan *data warehouse* untuk analisis data berskala besar. Sistem yang dibangun juga berpotensi dikembangkan lebih lanjut melalui integrasi orkestrator seperti Apache Airflow untuk otomatisasi *pipeline* serta penerapan model prediktif untuk mendukung analisis tren ketenagakerjaan digital secara lebih komprehensif.

DAFTAR PUSTAKA

- [1] A. Eweje and F. Ohaegbu, "Advances in Modern Data Stack Architectures for Scalable Data Integration and Business Intelligence," *Int. J. Multidiscip. Res. Growth Eval.*, vol. 2, no. 5, pp. 538–550, 2021, doi: 10.54660/IJMRGE.2021.2.5.538-550.
- [2] S. R. Cheruku, S. Jain, and A. Aggarwal, "Building Scalable Data Warehouses: Best Practices and Case Studies," *Darpan Int. Res. Anal.*, vol. 12, no. 01, pp. 80–99, 2024, doi: 10.36676/dira.v12.i1.87.
- [3] S. Pasupuleti, "The Medallion Architecture in Data Engineering : A Layered Approach to Data Quality and Governance," *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 62s, pp. 131–137, 2025, doi: 10.52783/jisem.v10i62s.13557.
- [4] R. Lakshmanasamy and G. Ganachari, "Integration of Dbt With Modern Data Stack Technologies," *Int. J. Multidiscip. Res.*, vol. 5, no. 5, pp. 1–6, 2023, doi: 10.36948/ijfmr.2023.v05i05.19750.
- [5] R. Lee, "layoffs.fyi." [Online]. Available: <https://layoffs.fyi/>
- [6] R. Vale, "Forecasting the 2022-23 tech layoffs using epidemiological models," pp. 1–8, 2023, doi: 10.48550/arXiv.2305.05210.
- [7] M. Armbrust, A. Ghodsi, R. Xin, and M. Zaharia, "Lakehouse : A New Generation of Open Platforms that Unify Data Warehousing and Advanced Analytics," in *11th Annual Conference on Innovative Data Systems Research (CIDR '21)*, CIDR Association, 2021. [Online]. Available: <https://www.semanticscholar.org/paper/Lakehouse%3A-A-New-Generation-of-Open-Platforms-that-Zaharia-Ghodsi/451cf5fc9786ed4f7e1d9877f08d00f8b1262121>
- [8] A. Safyan, M. U. Khan, M. Soomro, N. Sheikh, A. Rehman, and M. R. Tahir, "CLOUD-NATIVE DATA WAREHOUSING SOLUTIONS: ENHANCING SCALABILITY, SECURITY, AND PERFORMANCE IN BIG DATA ECOSYSTEMS," *Kashf J. Multidiscip. Res.*, vol. 09, no. 09, pp. 1–17, 2025, doi: 10.71146/kjmr603.
- [9] S. Salami, "Hub Star Modeling 2.0 for Medallion Architecture," 2025, doi: doi.org/10.48550/arXiv.2504.08788.

- [10] K. A. Hartono, M. I. Fianty, J. A. William, and J. C. Saputra, "Visualisasi Dinamika PHK Global dan Regional Menggunakan Tableau: Dampak Teknologi dan Ekonomi di Era Globalisasi," *J. Inform. dan Teknol.*, vol. 7, no. 2, pp. 521–532, 2024, doi: 10.29408/jit.v7i2.26487.
- [11] P. Hutabalian, B. R. P. Ginting, M. Q. Zaidan, N. Alfisyahrina, and C. Rozikin, "IMPLEMENTASI PIPELINE ETL/ELT DAN MODEL DIMENSIONAL UNTUK ANALISIS PENJUALAN SHOPEE MENGGUNAKAN POSTGRESQL, DOCKER, DAN APACHE SUPERTSET," *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 3, pp. 1321–1332, 2025, doi: 10.23960/jitet.v13i3S1.8093.
- [12] P. K. R. Gujjala, "Optimizing ETL Pipelines with Delta Lake and Medallion Architecture : A Scalable Approach for Large-Scale Data," *Int. J. Multidiscip. Res.*, vol. 6, no. 6, pp. 1–8, 2024, doi: 10.36948/ijfmr.2024.v06i06.55445.
- [13] L. Kaptsov, "Applying Postgis for Storage and Processing of Geospatial Data in Logistics System," *Am. J. Eng. Technol.*, vol. 07, no. 318, pp. 318–327, 2025, doi: 10.37547/tajet/Volume07Issue08-28.
- [14] N. R. Ni'mah, A. Larasati, C. N. Laksana, and E. Mohamad, "Business Intelligence System Design Based on Performance Monitoring Dashboard Using Online Analytical Processing (OLAP) Method," *J. Ilm. Tek. Ind.*, vol. 23, no. 2, pp. 191–204, 2024, doi: 10.23917/jiti.v23i02.6016.
- [15] S. Garudasu, I. Khan, and M. M. K. Dandu, "The Role of CI / CD Pipelines in Modern Data Engineering : Automating Deployments for Analytics and Data Science Teams," *IRE Journals*, vol. 5, no. 3, pp. 187–201, 2021, [Online]. Available: https://www.researchgate.net/publication/389435463_The_Role_of_CICD_Pipelines_in_Modern_Data_Engineering_Automating_Deployments_for_Analytics_and_Data_Science_Teams
- [16] A. Mayer, "BOOM AND BUST: JOB CHURN IN THE TECHNOLOGY INDUSTRY IN THE COVID-19 PANDEMIC ERA," 2023. [Online]. Available: https://scholarship.claremont.edu/scripps_theses/2046/
- [17] K. Rabuzin, M. Cerjan, and A. Lovrenčić, "Data Warehouse Design – Star Schema Synthesis Algorithm," *TEM J.*, vol. 14, no. 2, pp. 1707–1714, 2025, doi: <https://doi.org/10.18421/TEM142-68>.